

URIDINE ANALOGS AND TECHNIQUES FOR MAKING AND USING

BACKGROUND

5

Field of the Invention

This invention relates to uridine analogs and, in particular, to making and using nucleoside phosphates having a 5-aminouracil group.

10 Description of the Related Art

Uracil is a nucleic acid base, that, when bound to ribose or deoxyribose, forms uridine or deoxyuridine, respectively. Certain analogs of uridine and deoxyuridine have been prepared and used in different applications.

For example, deoxyribouridine analogs containing a 5-amino group have been prepared and used as anticancer, antiviral, or antibacterial agents. Such uses may be found, for example, in Shepard, Patent No. WO 2001/007088; Shepard, Patent No. WO 2001/007087; Shepard, *et al.*, Patent No. WO 99/37753; Shepard, Patent No. WO 99/08110; Wigerinck, *et al.*, *J. Med. Chem.*, 34:1767-1772 (1991); Cinatl, *et al.*, *Cancer Therapy*, 37:128-133 (1991); Beck, *et al.*, *Adv. Ext. Med. Biol.*, 195B:97-104 (1996); Desgranges, *et al.*, *Biochem. Pharmacol.*, 32:3583-3590 (1983); Balzarini, *et al.*, *Biochem. Pharmacol.*, 31:3673-3682 (1982); Hampton, *et al.*, *J. Med. Chem.*, 25:644-649 (1982); Balzarini, *et al.*, *Proc. Int. Roundtable Nucleosides, Nucleotides Biol. Appl.* 4:275-291 (1982); Hampton, *et al.*, *J. Med. Chem.*, 22:1524-1528 (1979); Hampton, *et al.*, *J. Med. Chem.*, 22:621-631 (1979); Lee, *et al.*, *Biochemistry*, 15:3686-3690 (1976); Leung, *et al.*, *Biochem. Med.*, 16:127-137 (1976); Cheng, *et al.*, *Biochemistry*, 13:1179-1185 (1974); or Roy-Burman, *et al.*, *Biochem. Pharmacol.* 19:2745-2756 (1970). Other nucleotide analogs have been described in, for example, Verdine, *et al.*, Patent No. WO 98/00433; Hayashibara, *et al.*, *J. Am. Chem. Soc.*, 113:5104-5106 (1991); Hayashibara, *et al.*, *Biochemistry*, 31:11265-11273 (1992); or Mascareñas, *et al.*, *J. Am. Chem. Soc.*, 115:373-374 (1993).

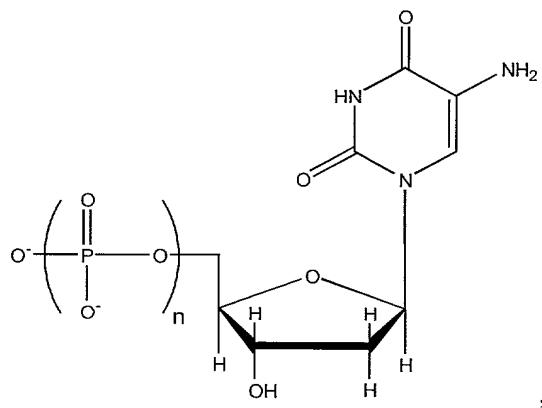
Other 5-amino deoxyribouridine analogs have been synthesized and analyzed, for example, in Ferrer, *et al.*, *Bioconjugate Chem.*, 8:757-761 (1997); Ferrer, *et al.*, *Nucleosides Nucleotides*, 15:907-921 (1996); Barawkar, *et al.*, *Bioorg. Med. Chem. Lett.*, 3:347-352

(1993); Eriksson, *et al.*, *Biochem. Biophys. Res. Commun.*, 176:586-592 (1991); Baerwolff, German Patent No. DD 137110; or *Tetrahedron*, 48:8505-8514 (1992). In some cases, 5-amino deoxyribouridine analogs have been derivitized with fluorescent compounds, for example, in Jadhav, *et al.*, *Nucleosides Nucleotides*, 16:107-114 (1997). 5-amino
5 deoxyribouridine analogs have also been prepared for DNA sequence determination, such as in Dyatkina, *et al.*, *Nucleic Acids Symp. Ser.*, 18:117-120 (1987).

SUMMARY OF THE INVENTION

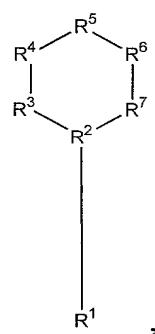
The present invention relates to nucleoside phosphates having a 5-aminouracil group.

10 In one aspect, the invention provides a composition. In one set of embodiments, the composition comprises a structure:



where n is at least 2.

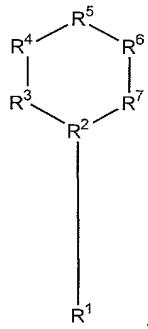
15 In another set of embodiments, the invention provides a composition comprising a structure:



where R¹ is capable of being integrated into a nucleic acid; R² is a carbon atom or a nitrogen atom; R³ comprises an atom; R⁴ and R⁵ are capable of participating in hydrogen-bonding to

adenine or an adenine analog; and at least one of R⁶ and R⁷ comprises an electron-rich substituent.

In another set of embodiments, the invention provides a composition comprising a structure:



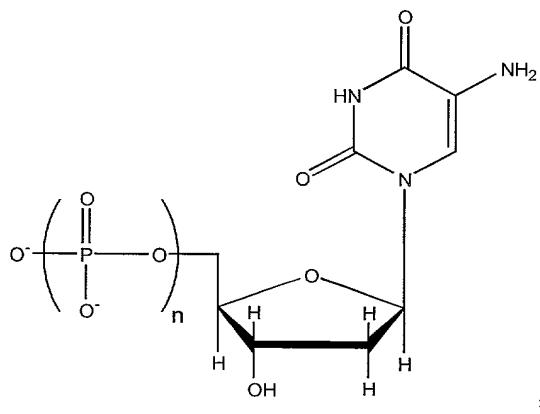
5

where R¹ is capable of being integrated into a nucleic acid; each of R², R³, R⁴, R⁵, R⁶ and R⁷ is a carbon atom or a nitrogen atom; at least one of R⁴ and R⁵ is capable of hydrogen bonding to or carries a substituent capable of hydrogen bonding to adenine or an adenine analog; and at least one of R⁶ or R⁷ carries an electron-rich substituent.

10 In each of the above embodiments, selection of R¹ through R⁷ is preferably made as described below in connection with the description of Fig. 1.

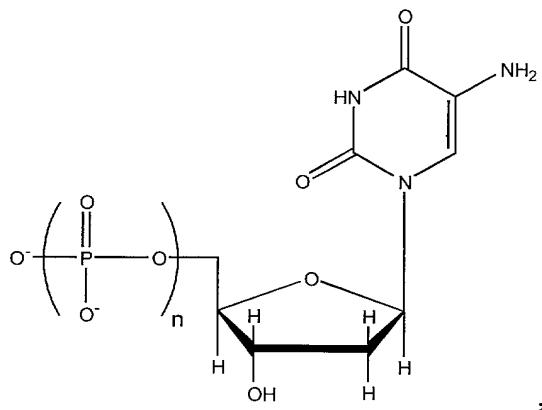
It is to be understood that any and all compositions disclosed herein can be used in connection with any and all methods, and vice versa.

15 In another aspect, the invention provides a method of synthesizing a compound. In one set of embodiments, the method comprises the steps of providing a species comprising a nucleoside, and reacting the species with a compound to form a composition comprising a structure:



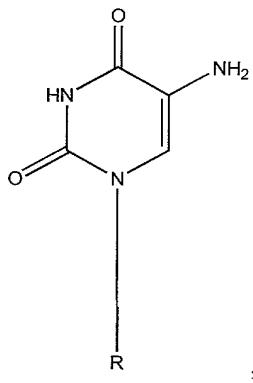
where n is at least 2.

In another set of embodiments, the invention provides a method of synthesizing a deoxyribonucleic acid. The method comprises the steps of providing a plurality of nucleotides, providing a composition comprising a structure:



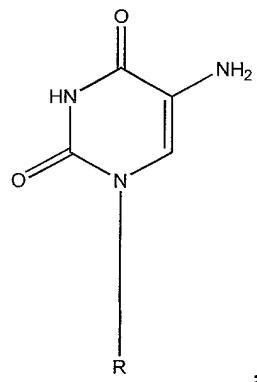
5 where n is at least 2, and reacting the nucleotides and the composition to produce a deoxyribonucleic acid.

In another set of embodiments, the invention provides a method of synthesizing a deoxyribonucleic acid. The method comprises the steps of providing a nucleic acid comprising a base at a position within the nucleic acid, and synthesizing a deoxyribonucleic acid using the nucleic acid as a template. The base in the nucleic acid is one of thymine or uracil. The base at a position within the deoxyribonucleic acid corresponding to the position of the base within the nucleic acid is substituted by an unnatural base comprising a structure:



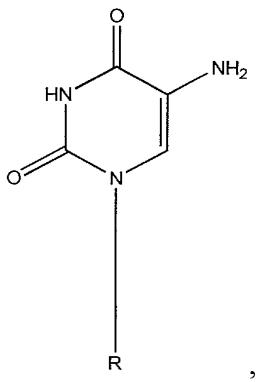
where R is a covalent bond.

15 In another set of embodiments, the invention provides a method of analyzing a deoxyribonucleic acid. The method comprises the steps of providing a deoxyribonucleic acid comprising an unnatural base, and determining a position of the unnatural base within the deoxyribonucleic acid. The unnatural base comprises a structure:



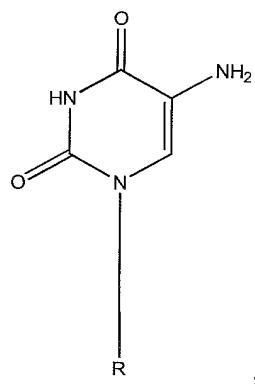
where R is a covalent bond.

In another set of embodiments, the invention provides a method of characterizing a compound. The method comprises the steps of providing a compound, contacting the compound with a deoxyribonucleic acid comprising an unnatural base, and determining a degree of binding between the compound and the deoxyribonucleic acid. The unnatural base comprises a structure:



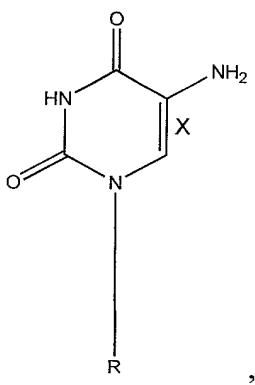
where R is a covalent bond.

In another set of embodiments, the invention provides a method of oxidizing a deoxyribonucleic acid. The method comprises the steps of providing a deoxyribonucleic acid comprising an unnatural base, and reacting the deoxyribonucleic acid with an oxidizing agent. The unnatural base comprises a structure:



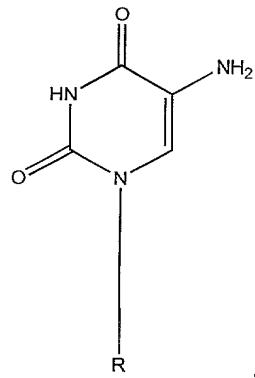
where R is a covalent bond.

In another set of embodiments, the invention provides a method of oxidizing a deoxyribonucleic acid. The method comprises the step of providing a deoxyribonucleic acid comprising an unnatural base comprising a structure:



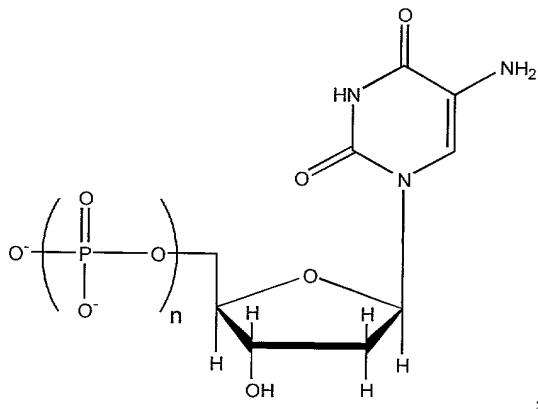
where R is a covalent bond. The method also comprises the step of reacting the deoxyribonucleic acid with a compound to cleave double bond X.

In another set of embodiments, the invention provides a method of oxidizing a deoxyribonucleic acid. The method comprises the steps of providing a deoxyribonucleic acid comprising an unnatural base at a position within the deoxyribonucleic acid, and reacting the deoxyribonucleic acid with a compound to cleave the deoxyribonucleic acid at the position of the unnatural base. The unnatural base comprises a structure:



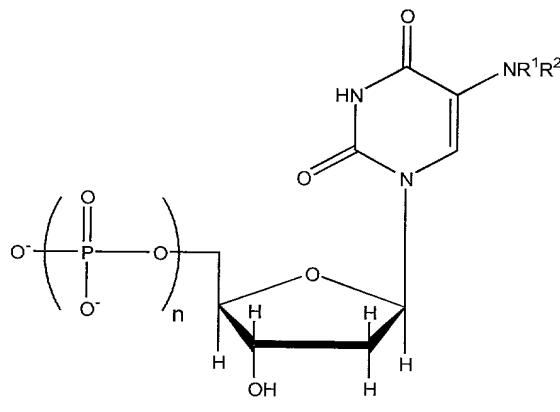
where R is a covalent bond.

In another set of embodiments, the invention provides a method of synthesizing an amine. The method comprises the step of providing a composition comprising a structure:



5

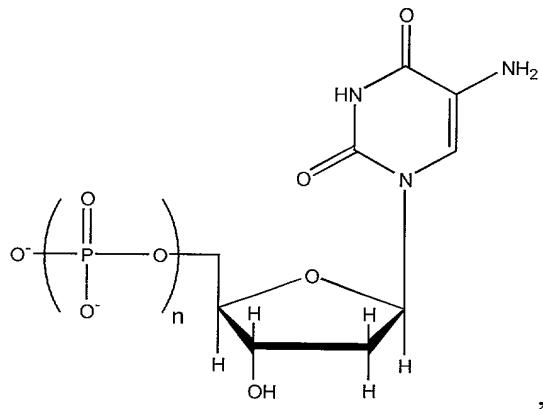
where n is at least 2. The method also comprises the step of reacting the NH₂ group of the composition with a compound to form an amine comprising a structure:



where at least one of R¹ and R² comprises a carbon atom.

10 In another set of embodiments, the invention provides a method of synthesizing a compound. The method comprises the step of providing a composition comprising a

structure:



where n is at least 2, and coupling a reporting group to the compound.

Other advantages, novel features, and objects of the invention will become apparent
5 from the following detailed description of the invention when considered in conjunction with
the accompanying drawings, which are schematic and which are not intended to be drawn to
scale. In the figures, each identical, or substantially similar component that is illustrated in
various figures is represented by a single numeral or notation. For purposes of clarity, not
every component is labeled in every figure, nor is every component of each embodiment of
10 the invention shown where illustration is not necessary to allow those of ordinary skill in the
art to understand the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Non-limiting embodiments of the present invention will be described by way of
15 example with reference to the accompanying drawings in which:

FIG. 1 is a structural diagram of one embodiment of the invention;

FIG. 2 is a reaction pathway of one embodiment of the invention;

FIG. 3 is a structural diagram, showing how one embodiment of the invention may
hydrogen-bind to an adenine base; and

20 FIG. 4 is a structural diagram, showing how another embodiment of the invention
may hydrogen-bind to an adenine base.

DETAILED DESCRIPTION

Uridine analogs and techniques for making and using uridine analogs are provided by
25 the present invention. They may be used in many ways, including measuring chemical

interactions between nucleic acids and other compounds, or sequencing nucleic acids. These uridine analogs may include nucleoside phosphates having a 5-aminouracil group. These nucleotides may be incorporated into a nucleic acid as an unnatural base, as a substitute for uridine or thymine. The nucleic acid may then be treated with an oxidizing agent and an
5 alkaline solution, which may cause cleavage of the nucleic acid at the position of the unnatural base. Additional compounds may also be derivitized onto the amino group, allowing other functionalities to be added to the nucleoside phosphate, or to the nucleic acid incorporating the nucleoside phosphate.

As used herein, the term “base” is given its ordinary meaning as is understood in
10 biochemistry and similar fields of study, referring to nucleic acids and similar compounds. Chemical bases such as sodium hydroxide will be referred to as “alkaline compounds,” “alkaline materials,” and the like.

“Nucleotides” and “nucleosides” are given their ordinary meanings as is understood in the field of biochemistry. A nucleoside generally includes a base and a sugar, while a
15 nucleotide generally includes a base, a sugar, and at least one phosphate. The atoms within the base and the sugar are conventionally numbered, with primes typically indicating the atoms of the sugar group. A “natural base” may be any one of the commonly observed bases found on nucleic acids such as deoxyribonucleic acid (“DNA”) or ribonucleic acid (“RNA”). These may include adenine, thymine, guanine, cytosine, or uracil. Correspondingly, an
20 “unnatural base” is a base that is not commonly observed bases found on nucleic acids such as DNA or RNA.

A “deoxyribo-” generally indicates a ribose sugar where one of the hydroxide groups located on the ribose ring has been substituted by hydrogen. Generally, this is the 2’ hydroxide group.

25 As used herein, “hydrocarbon,” “alkyl,” and similar terms includes not only alkyl groups, but also alkenyl, alkynyl, cycloalkyl, aryl, alkaryl, aralkyl, and the like. Examples of such hydrocarbon groups may include methyl, propenyl, ethynyl, cyclohexyl, phenyl, tolyl, benzyl, hydroxyethyl and the like. Hetero groups may include -O-, -CONH-, -CONHCO-, -NH-, -CSNH-, -CO-, -CS-, -S-, -SO-, -(OCH₂CH₂)_n- (where n may range from 1 to 10), -
30 (CF₂)_n- (where n may range from 1 to 10), olefins, and the like. “Hydrocarbon,” “alkyl,” and similar terms are also meant to include alcohols and hydrogen. By way of example, the “alkyl” may comprise an alkyl group, preferably having 1 to 24 carbon atoms, more

preferably 1 to 18 carbon atoms; an alkenyl group, preferably having 2 to 4 carbon atoms; an alkylamino group, preferably having 1 to 8 carbon atoms, and optionally substituted on the nitrogen atom with one or, preferably two alkyl groups, preferably having 1 to 4 carbon atoms; an alkyl group, preferably having 1 to 4 carbon atoms, having a five- or six-membered 5 heterocyclic ring as a substituent; an allyloxyalkyl group, preferably having up to 12 carbon atoms; an alkoxyalkyl group, preferably having a total of 2 to 12 carbon atoms; an aryloxyalkyl group, preferably having 7 to 12 carbon atoms; an aralkyl group, or the like.

As used herein, "R" may indicate not only single atoms, but groups of atoms which may be connected by any suitable bond, for example, ionic, hydrogen, van der Waals, or 10 hydrophobic bonds. Preferably, R indicates an atom, or a group of atoms connected via covalent bonds.

The term "determining" refers to quantitative or qualitative analysis of a species via, for example, spectroscopy, radioactive counting, HPLC analysis, nuclear magnetic resonance, and the like. "Determining" also refers to detecting or quantifying interactions between 15 species, e. g. detection of the binding between two species.

"Derivative," "analog," "chemical derivative," "derivatizing," and similar terms are given their ordinary meanings as well-known in the fields of chemistry, biochemistry, or biology. A derivative may be any chemical substance structurally related to another chemical substance and at least theoretically derivable from it. An analog may be a chemical or 20 biological species that is similar enough to a parent species that it may substitute for the parent species in at least one set of chemical or biochemical interactions. For example, if the parent compound is a nucleic acid base, the analog may be able to substitute for the nucleic acid base, and be able to participate in reactions involving the nucleic acid base, such as hydrogen-bonding with another nucleic acid, or interactions with proteins or enzymes, such 25 as a ribosome.

A "dalton" (Da) is an alternate name for the unified atomic mass unit (grams/mole) that is well known in biochemistry and allied fields. The dalton is accepted by SI as an alternate name for the unified atomic mass unit. Similarly, a "kilodalton" (kDa) is 1000 daltons.

30 "Small molecule," as used herein, typically means a molecule less than 5 kilodalton, more typically less than 1 kilodalton.

“Proteins” and “peptides” are well-known terms in the art, and are not precisely defined in the art in terms of the number of amino acids that each includes. As used herein, these terms are given their ordinary meaning in the art. Generally, peptides are amino acid sequences of less than about 100 amino acids in length, but can include sequences of up to 5 300 amino acids. Proteins generally are considered to be molecules of at least 100 amino acids, but may be less in some cases.

A molecule “capable of integrating” into a second molecule generally refers to a molecule that can be integrated or otherwise attached to a second molecule, without requiring the use of further synthetic chemical reactions that modify the structure of the molecule into a 10 different, stable structure, prior to introduction of the molecule into the second molecule.

Fig. 1 illustrates 5-amino-2'-deoxyuridine diphosphate, in accordance with one embodiment of the invention. In this figure, modified uridine nucleotide 1 comprises a uridine moiety 2 having a 5-amino group 3, a 2'-deoxyribose sugar 4, and two phosphate groups 5, 6. As shown in Fig. 1, the nucleotide has two phosphate groups. However, in other 15 embodiments of the invention, additional phosphate groups may also be present or the phosphate groups may be located in different positions. The phosphate groups may be located anywhere on the molecule, attached to each other or to different portions of the molecule, for example, on the 2'-carbon of the deoxyribose ring. In some embodiments, the phosphates may be attached to the 5'-end of the sugar, and arranged in a linear fashion.

The sugar of the nucleotide is shown in Fig. 1 as being deoxyribose. However, the 20 sugar may be any sugar that can be incorporated into a nucleic acid. For example, other substituents may be present on the sugar, or the sugar may have a different number of atoms within the ring. For example, one or more fluorine atoms or hydroxide groups may be connected to the 2'- or 3'-carbon atoms of the sugar. Alternatively, the phosphate or the base 25 may be connected to the sugar by an α linkage instead of a β linkage.

As shown in Fig. 1, the base is a uracil moiety having an attached amino group on the 5-carbon atom. However, the base may also be another pyrimidine, such as thymine. Any 30 electron-rich group may be substituted onto the pyrimidine in any position, such as on the 5-carbon or 6-carbon atoms, for example, but not limited to, a hydroxide, a halogen such as fluorine or chlorine, an alkyl, an amino group, an alkylamino group, or a dialkylamino group. Adding an electron-rich substituent to the base may provide better stability in certain cases, for example, during elevated temperatures, such as may be encountered in a polymerase

chain reaction device. The base may also have other substituents. For example, additional electron-rich substituents may be attached to the base to, for example, enhance stability of the molecule. As an example, a fluorine or a chlorine atom may be attached to the 6-carbon atom of a 5-aminouracil moiety. Alternatively, one or more of the oxygen atoms may be modified
5 or replaced by another type of atom, for example, sulfur. Combinations of these substituents may also be envisioned.

It is to be understood that the modified nucleotides of the invention may have a wide range of variations. For example, other bases or sugars may be incorporated into the nucleotide. In other cases, other substituent groups may be added to the base or the sugar, for
10 example, but not limited to, electron-rich, reactive or enzymatic moieties. Other moieties may include fluorescent moieties, such as fluorescein, or radioactive molecules, such as 14 carbon or 35 sulfur. Alternatively, a reporter molecule may be derivitized onto the nucleotide. For example, a protein, a small molecule such as biotin, or a catalyst such as green fluorescent protein (“GFP”) may be attached to the nucleotide.
15

The above and other modifications may confer certain properties on the modified nucleotides of the invention. The modified nucleotides of the invention, in some cases, may be analogous to certain nucleic acid bases, such as uridine or thymine. In some cases, the analog has a structure, or an electronic configuration, similar to that of a natural nucleic acid base, and may not cause distortion of the nucleic acid strand or alteration of the secondary
20 structure of the nucleic acid strand when introduced into the nucleic acid strand. In other cases, the modified nucleotides of the invention may participate in hydrogen-bonding to another nucleotide, such as hydrogen-bonding to an adenine base or an adenine analog. The 3-hydrogen and 4-oxygen atoms of the modified nucleotide may be able to hydrogen-bond to another nucleotide. For example, in Fig. 3, nucleotide 1 may hydrogen-bind to the 1-nitrogen
25 atom or the 6-amino group of adenine base 10 through hydrogen bonds 11 and 12. Fig. 4 shows an adenine base 10 hydrogen bonding to groups R⁴ and R⁵ on molecule 20 through hydrogen bonds 21 and 22, respectively. Other hydrogen-bonding configurations may also be possible.

The modified nucleotides may also be targeted and cleaved as desired. In some cases,
30 modified nucleotides which have been integrated into a nucleic acid strand may be cleaved without altering the rest of the nucleic acid strand. The modified nucleotides of the invention may be used, in some cases, to study the interaction between a nucleic acid and another

molecule, or to study the structure of the nucleic acid. In other cases, the modified nucleotides of the invention may be used to sequence a nucleic acid. The modified nucleotides of the invention may also be used to expand the abilities of nucleic acids, for example, to introduce various chemical functionalities into a nucleic acid strand, without
5 distorting its native confirmation.

The modified nucleotides of the invention may be synthesized and prepared using any suitable technique. For example, the modified nucleotides may be synthesized from deoxyuridine or a substituted deoxyuridine, such as 5-chloro-2'-deoxyuridine or 5-bromo-2'-deoxyuridine. The 5-amino group may be added to a substituted deoxyuridine to produce an
10 aminated deoxyuridine using any suitable method, such as amination with an excess of ammonia under appropriate conditions. Alternatively, starting with a nitrated substituent, such as 5-nitro-2'-deoxyuridine, hydrogenation of the starting material over a palladium catalyst under hydrogen gas may also yield aminated deoxyuridines. Phosphorylation of aminated deoxyuridines may occur using any suitable technique, for example, reaction with a
15 pyrophosphate, such as tributylammonium pyrophosphate. The number of deoxyuridine phosphorylation events may be controlled by any appropriate method, such as limiting the concentration of pyrophosphates or altering the reaction conditions.

It should be understood that the modified nucleotides of the present invention are not limited to only those nucleotides that can be produced using any of the above reactions.
20 Alternative reaction pathways to produce 5-amino-2'-deoxyuridine or any of the other modified nucleotides of the present invention may also be contemplated. For instance, alternative phosphorylation or amination pathways may be envisioned, or the phosphorylation and amination steps may be reversed. Additional reaction pathways, for example, to add protecting or deprotecting groups or to add additional substituents, for
25 example, to the uridine ring or to the amino group, may also be considered. Biologically-derived synthesis of the nucleotides may also be possible, for example, through the use of genetic engineering, such as through the use of various post-translational modification pathways.

The modified nucleotides may be added to a nucleic acid by any suitable technique.
30 Examples of suitable techniques include polymerase chain reactions, solid-phase synthesis techniques, for example, with the use of a dimethoxytrityl protecting groups, or genetically engineering organisms, such as bacteria or mammalian cells. Dideoxy derivatives of the

modified nucleotides may also be useful in certain forms of DNA synthesis. Polymerase chain reactions (“PCR”), for example, as described in Mullis et al., U. S. Patent No. 4,965,188 and incorporated herein by reference, may be particularly suitable for certain embodiments of the invention. One particular set of PCR conditions that may be useful
5 include heating the reaction mixture including the nucleic acid to 94 °C for 3 minutes, then repeating the following procedure for a predetermined number of times, for example, 29 times: heating the mixture to 94 °C for 1 minute, cooling the mixture to 55 °C for 90 seconds, followed by warming the mixture to 72 °C for 90 seconds. A final annealing step of warming the mixture at 72 °C for 7 minutes may also be added. The modified nucleotides may be
10 added to any position within the nucleic acid, and additional nucleotides (modified or unmodified) may further be added to the growing nucleic acid strand after modified nucleotides have been added. Other techniques for adding the modified nucleotide to a nucleic acid strand may be contemplated without departing from the scope of the invention, for example, using genetic engineering techniques, either *in vitro* or *in vivo*, such as within a genetically modified bacterium.
15

Targeted cleavage of the modified nucleotides of the invention integrated in a nucleic acid may be useful in some cases, for example, to determine the structure or the composition of the nucleic acid. The modified nucleotides may be cleaved using any suitable technique, such as enzymatic cleavage using a restriction endonuclease. Some nucleotides may be
20 particularly susceptible to cleavage by exposure to an oxidative agent and an alkaline compound. Exposure of the modified base to an oxidative agent followed by an alkaline compound may destroy or modify the base. An illustration of the reaction mechanism by which the modified nucleotide is cleaved is shown in Fig. 2. The oxidative agent may first attack the base near the electron-poor end of the pyrimidine ring, which may cause the
25 pyrimidine ring to open. After the pyrimidine ring has been opened, exposure to an alkaline compound may cause the nucleotide to degrade. Any oxidative agent which may attack the modified nucleotide may be used, for example, potassium permanganate, hydrogen peroxide, potassium dichromate, aqueous periodates, or aqueous iodine. Similarly, any alkaline compound which may assist in cleavage of the modified base may be used; suitable alkaline
30 compounds may include alkali metal hydroxides, such as sodium hydroxide or potassium hydroxide, or organic alkaline materials, such as piperidine, or pyrrolidine. There may be certain optimal concentrations of these alkaline compounds. For example, particularly useful

concentrations of sodium hydroxide may be between 0.1 to 0.5 M preferably between 0.2 and 0.4 M. A suitable concentration of an organic alkaline may be between 5 and 15%, preferably between 8 and 12%. If the modified base is within a nucleic acid, the cleavage of the base by the oxidative agent and the alkaline compound may split the nucleic acid into two
5 fragments, as shown in Fig. 2. In certain embodiments of the invention, this reaction may be used in sequencing the nucleic acid, or in determining how the nucleic acid interacts with proteins or other molecules, for example, as described below.

In one set of embodiments of the invention, the modified nucleotides may be used to measure or analyze interactions between nucleic acids and other molecules. The molecular
10 interactions may be measured by any suitable technique, for example, DNA footprinting, as described in, for example, Larson *et al.*, *Bioorganic Chemistry: Nucleic Acids*, Hecht, Ed., Oxford University Press, New York, 324-342 (1996). The other molecules may include proteins, such as proteins that bind noncovalently to DNA or RNA, smaller peptides, other nucleic acids, hormones, intercalating agents, or small molecules that may interact with
15 nucleic acids. The small molecules typically have a molecular weight of less than about 2000 daltons, preferably less than about 1000 daltons. The modified nucleotides of the invention may be suitable for detecting and measuring various types of noncovalent interactions between the nucleic acid and other molecules, for example, but not limited to, ionic, van der Waals, entropic, steric, hydrogen-bond, or hydrophobic interactions and the like. For
20 example, by substituting a modified nucleotide for a thymine or a uracil base on a nucleic acid, an interaction between the nucleic acid and another molecule may be altered in some fashion. Measurement of the change in the interaction may be useful in determining the degree of binding between the nucleic acid and the molecule, or the stereospecificity of the two molecules. Other molecular interactions may also be analyzed in this fashion, for
25 example, but not limited to, DNAzyme/substrate, promoter/DNA, DNA/RNA, DNA/DNA, DNA/protein, DNA/transcription factor, DNA/enzyme, DNA/ribosome, DNA/ribozyme, and the like.

The modified nucleotides of the invention may also be used to determine the structure or the composition of a nucleic acid, such as DNA or RNA. Any technique may be used to
30 determine the structure or the composition, for example, X-ray crystallography. The nucleotides of the invention may be used to sequence a nucleic acid. A sequence of nucleic acids incorporating the modified nucleotides of the invention may still continue to be grown

by any suitable technique, as the modified nucleotides may not interfere or sterically hinder nucleic acids synthesis enzymes, such as polymerase. Thus, nucleic acid sequences may continue to be synthesized after incorporation of the modified nucleotides. As the modified nucleotides may not interfere with DNA synthesis, the distribution of the modified
5 nucleotides within the growing nucleic acid strand can be controlled and may be distributed evenly within the nucleic acid strand.

The modified nucleotides incorporated into a nucleic acid may later be cleaved using a combination of a oxidative agent and any alkaline compound, as previously described. The resulting nucleic acid fragments may then be individually sequenced by any suitable
10 technique. Comparison of the fragments may then be used to reconstruct the original nucleic acid sequence. As specific individual nucleotides may be targeted for cleavage using this technique, specific cleavage and sequencing of nucleic acids may be performed. The nucleic acids may be characterized by any suitable method of detecting or sequencing the nucleic acids. Suitable techniques may include, for example, but not limited to, capillary gel
15 electrophoresis, radioactive labeling, fluorescent labeling, or other gel sequencing techniques.

In another set of embodiments of the invention, the modified nucleotides may be integrated within nucleic acids to expand the abilities of the nucleic acids to perform chemically. As a non-limiting example, nucleic acids having the modified bases of the invention may be used to produce novel ribozymes or DNAzymes. The novel ribozymes or
20 DNAzymes may be designed, for example, by using directed evolution or by using computational design techniques, such as *ab initio* calculations. In other embodiments of the invention, the modified nucleotide may further be modified with, for example, additional functional groups, proteins, enzymes, fluorescent moieties, or radioactive atoms. The functional groups may be any functional group, such as an alkyl group. Additionally, two or
25 more groups may be attached to the same base, providing additional functionalities. Further modifications may also be possible. By adding additional chemical groups to the modified nucleotide, additional properties or characteristics for the nucleotide or for the molecule incorporating the modified nucleotide may be provided in this fashion.

The function and advantages of these and other embodiments of the present invention
30 will be more fully understood from the following examples. These examples are intended to be illustrative in nature and are not considered to be limiting the scope of the invention.

Example 1: Synthesis of 5-amino-2'-deoxyuridine

This example demonstrates the synthesis of 5-amino-2'-deoxyuridine in one embodiment of the invention.

All reactions were run under argon unless otherwise indicated. 5-bromo-2'-

5 deoxyuridine (3.0 g) was loaded into a steel bomb. The bomb was cooled to -78 °C and liquid ammonia (20 ml) was carefully added. The starting material dissolved readily. The bomb was then sealed and stirred at 55-60 °C. After 44 hours, the bomb valve was released and ammonia allowed to boil off. Particular attention was paid to avoid vigorous boiling of ammonia to avoid loss of product. Precoated silica-gel plates were used to perform thin-layer chromatography ("TLC"). Detection was by exposure to ultraviolet light, or dipping the plates into 5% sulfuric acid in methanol, followed by heating to 200 °C. TLC analysis using a 2:8 ratio of methanol to methylene chloride showed that the reaction was complete and only one product was formed.

10 Pyridine (25 ml) was added to the brown oil in the bomb and the resulting suspension was transferred to a 100 ml beaker. The bomb was washed with 3 ml pyridine, which were then added to the same beaker. Gravity filtration of this suspension afforded a white precipitate which was found to be ammonium bromide (NH₄Br).

15 The filtrate was concentrated *in vacuo* to an oil, which was diluted with methanol (10 ml) and added to silicon dioxide (20 ml). The resulting slurry was concentrated to dryness *in vacuo* and loaded onto a 3 cm diameter chromatographic column. The product was eluted with a methanol/methylene chloride solvent system (starting with 5% methanol in methylene chloride, and increasing linearly to 15% methanol). Appropriate fractions were pooled together and concentrated *in vacuo*. A white solid (1.54 g, 65% yield) was isolated and characterized.

20 Nuclear magnetic resonance ("NMR") was recorded at 500 MHz for ¹H-NMR recorded and 400 MHz for ¹³C-NMR. NMR analysis of the white solid is as follows:

R_f(MeOH:DCM = 2:8) 0.26, UV: λ_{max} = 294 nm, ε = 7400. ¹H-NMR (CD₃OD): 7.22 ppm (s, 1H), 6.28 ppm (t, 1H), 4.33 ppm (dd, 1H), 3.83 ppm (ddd, 1H), 3.69 ppm (dd, 2H), 2.15 ppm (ddd, 2H). ¹H-NMR(⁶d-DMSO): 11.35 ppm (s, 1H), 6.92 ppm (s, 1H), 6.19 ppm (s, 1H), 5.21 ppm (s, 1H), 4.91 ppm (s, 1H), 4.18 ppm (dd, 1H), 4.14 ppm (s, 1H), 3.72 ppm (ddd, 1H), 3.42 ppm (dd, 2H), 3.18 (s, 1H), 1.95 ppm (ddd, 2H).

Thus, this example demonstrates the synthesis of 5-amino-2'-deoxyuridine, according to one embodiment of the invention.

Example 2: Synthesis of 5-amino-2'-deoxyuridine-5'-triphosphate

5 This example illustrates the synthesis of 5-amino-2'-deoxyuridine-5'-triphosphate from 5-amino-2'-deoxyuridine, according to an embodiment of the invention.

10 5-amino-2'-deoxyuridine and anhydrous tributylammonium pyrophosphate were thoroughly dried before use by storing them in a desiccator charged with P₄O₁₀ under high vacuum for 3 days. Tributylammonium pyrophosphate and PO(CH₃O)₃ were dried on CaH₂ overnight and then distilled at reduced pressure (house vacuum).

15 5-amino-2'-deoxyuridine (330 mg) was dissolved in PO(CH₃O)₃ (2.7 mL) in a flame dried 25 ml flask. The resulting pale-yellow solution was cooled to -15 °C and POCl₃ (190 µl) was slowly added. The contents of the flask were stirred at -10 °C for 3 hours with periodical monitoring of reaction progress by TLC, using procedures similar to those described above.

20 After 3 hours, the flask was subject to high vacuum for 15 minutes, keeping the flask at 0 °C, to remove unreacted POCl₃. A solution of tributylammonium pyrophosphate (1.24 g) in dry dimethylformamide (13.58 mL) was added to the reaction mixture, and the mixture was stirred for 2 minutes at 0 °C. The mixture was then quenched with 1.5 M tetraethylammonium bromide (20 ml, pH = 7.5), and further stirred at room temperature for 30 minutes.

25 The reaction mixture was lyophilized three times, then purified by ion-exchange chromatography using linear gradient of 1.5 M tetraethylammonium bromide. The ion exchange chromatography settings included a flow rate of 5 ml/min over diethylaminoethyl (“DEAE”) SEPHADEX™ A-25 beads (Pharmacia Corp., Peapack, NJ), with detection at 254 nm.

30 Part of the isolated triphosphate was then lyophilized 5 times and further purified by HPLC. HPLC conditions included a flowrate of 2 ml/min over ODS ULTRASPHERE™ (10 x 250 mm, Beckman Coulter, Inc., Fullerton, CA), with detection at 294 nm. The solvent system included 100 mM triethylamine acetate (“TEAA”) and CH₃CN in the following proportions: 1.5% CH₃CN (0 to 10 min), linear gradient to 4% CH₃CN (10 to 40 min), linear gradient to 10% CH₃CN (40 to 41 min), 10% CH₃CN (41 to 55 min), linear gradient to 1.5%

CH₃CN (55 to 56 min) and 1.5% CH₃CN (56 to 65 min). Isolated yielded 25 mg of HPLC-purified, hygroscopic white crystals.

Analytical HPLC conditions included a flowrate of 1 ml/min with detection at 294 nm over an ODS ULTRASPHERE™ 5 μm C₁₈ column (4.6 x 250 mm). The solvent system

5 used 100 mM TEAA (pH = 7.0) with a linear gradient (0 to 15% CH₃CN) over 25 min.

Analysis of the product was as follows: R_f(isopropanol:water:ammonia = 6:2:2) 0.16, with an analytical HPLC retention time of 9.7 minutes. ¹H-NMR (NH₄HCO₃/(NH₄)₂CO₃ buffer, pH = 8.5 in D₂O): 7.46 ppm (s, 1H), 6.36 ppm (t, 1H), 4.21 ppm (dd, 1H), 4.17 ppm (dd, 2H), 2.37 ppm (ddd, 2H). ³¹P-NMR (NH₄HCO₃/(NH₄)₂CO₃ buffer, pH = 8.5 in D₂O) : -9.7

10 ppm (d), -10.45 ppm (d), -22.3 ppm (t).

While several embodiments of the invention have been described and illustrated herein, those of ordinary skill in the art will readily envision a variety of other means and structures for performing the functions and/or obtaining the results or advantages described herein, and each of such variations or modifications is deemed to be within the scope of the present invention. More generally, those skilled in the art would readily appreciate that all parameters, dimensions, materials, and configurations described herein are meant to be exemplary and that actual parameters, dimensions, materials, and configurations will depend upon specific applications for which the teachings of the present invention are used. Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. It is, therefore, to be understood that the foregoing embodiments are presented by way of example only and that, within the scope of the appended claims and equivalents thereto, the invention may be practiced otherwise than as specifically described. The present invention is directed to each individual feature, system, material and/or method described herein. In addition, any combination of two or more such features, systems, materials and/or methods, if such features, systems, materials and/or methods are not mutually inconsistent, is included within the scope of the present invention. In the claims, all transitional phrases or phrases of inclusion, such as "comprising," "including," "carrying," "having," "containing," and the like are to be understood to be open-ended, i.e. to mean "including but not limited to." Only the transitional phrases or phrases of inclusion "consisting of" and "consisting essentially of" are to be interpreted as closed or semi-closed phrases, respectively.

What is claimed is: